PROCEEDINGS OF TDWG

Biodiversity Information Science and Standards 1: e20214
doi: 10.3897/tdwgproceedings.1.20214

OPEN ACCESS

Conference Abstract

Global Biotic Interactions: A Catalyst For Integrating Existing Interaction Datasets, Connecting Data Curators And Developing Data Exchange Methods

Jorrit H Poelen ‡

‡ 400 Perkins St Apt 104, Oakland, CA, United States of America

Corresponding author: Jorrit H Poelen (jhpoelen@xs4all.nl)

Received: 11 Aug 2017 | Published: 11 Aug 2017

Citation: Poelen J (2017) Global Biotic Interactions: A Catalyst For Integrating Existing Interaction Datasets, Connecting Data Curators And Developing Data Exchange Methods. Proceedings of TDWG 1: e20214. https://doi.org/10.3897/tdwgproceedings.1.20214

Abstract

Since 2013, Global Biotic Interactions (GloBI, globalbioticinteractions.org, Poelen et al. 2014) has taken an opportunistic, decentralized approach to integrating, and make accessible, existing species interaction datasets. Rather than expecting dataset curators to conform to some publication regime, methods were developed to automatically and algorithmically discover, parse and link existing datasets without the need to reformat, relocate, or transfer ownership of, the existing dataset. The automated nature of GloBI helps to: (a) automate propagation of dataset updates (b) quickly detect data integration issues (e.g. outage, change in data format), (c) integrate new datasets without having to contact some central office, (d) avoid permanent data loss due to software integration bugs, and, last but not least, (e) access to datasets even after GloBI goes away.

As far back as 1927, Charles Elton, a founding father of modern ecology, realized the importance of linking natural history knowledge stored in professional journals while acknowledging the value of local (amateur) knowledge. Despite technological advances, details on how species interact are only still largely available by studying professional journals, manually inspecting datasets or striking up a conversation with a ecologist, farmer or citizen scientist. The lack of access to species interaction data is known as the Eltonian shortfall (Hortal et al. 2015). GloBI's mission is to address this shortcoming.

© Poelen J. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

By borrowing from software engineering practices such as test driven development and continuous integration, re-purposing freely available platforms such as GitHub, Zenodo, Travis CI and integrating with many existing biodiversity services (e.g. globalnames.org, eol.org, crossref.org, geonames.org), GloBI has grown to include about 2.8M interaction records spanning 100k taxa (source: globalbioticinteractions.org/references, 17 July 2017) and has established bi-directional links to projects including, but not limited to, the NCBI Taxonomy, World Register of Marine Species, Encyclopedia of Life and iNaturalist.

As GloBI continues to link existing species interaction datasets, and form a loosely affiliated community of data curators, educators and (citizen) scientists, the data integration platform is well-suited to play an active and experimental role in the development of novel methods to more easily mobilize and integrate species interaction data in an effort to realize Charles Elton's dream to "[...] *provide conceptions which can link up into some complete scheme the colossal store of facts about natural history which has accumulated up to date in this rather haphazard manner.* [...]" (Elton 1927).

Keywords

ecology, species interactions, ecological informatics, species associations, data integration

Presenting author

Jorrit H. Poelen

References

- Elton C (1927) Animal ecology, by Charles Elton; with an introduction by Julian S. Huxley. Macmillian Compan https://doi.org/10.5962/bhl.title.7435
- Hortal J, Bello Fd, F. Diniz-Filho JA, Lewinsohn T, Lobo J, Ladle R (2015) Seven Shortfalls that Beset Large-Scale Knowledge of Biodiversity. Annual Review of Ecology, Evolution, and Systematics 46 (1): 523-549. https://doi.org/10.1146/annurev-ecolsys-112414-054400
- Poelen J, Simons J, Mungall C (2014) Global biotic interactions: An open infrastructure to share and analyze species-interaction datasets. Ecological Informatics 24: 148-159. https://doi.org/10.1016/j.ecoinf.2014.08.005